Lecture 1: Introduction to mathematical optimization

Rajat Mittal *

IIT Kanpur

1 Mathematical optimization

Most of the problems in this world are optimization. You have to maximize (happiness/peace/money) or minimize (poverty, grief, wars etc.). Unfortunately we are not solving any of those problems. On a smaller scale optimizing time in the production cycle of some industry, optimizing tax in a tax-return, optimizing length in a tour etc. are mathematical optimization problems we encounter in our daily life.

Formally, any problem of the form:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq b_i \ i = 1, 2, \cdots, m \end{aligned}$$

is called an optimization problem. Here f_o is the objective/optimization function and $f_i \leq b_i$ are called constraints. The task here is to find the max/min value of $f_0(x)$, s.t., x satisfies all the constraints. The set of all x which satisfy all the constraints is called the *feasible region.*

$$S = \{x : f_i(x) \leq b_i \quad i = 1, 2, \cdots, m\}$$

A vector $x*$is called "optimal"/solution if it has the smallest objective value among all vectors which satisfy the constraints. So for any z ,s.t., $f_i(z) \leq b_i$, we know, $f_0(z) \geq f_0(x*)$.

- Take our/mine favorite example of world peace, x-actions, f_0- peace function, f_i- many, dont kill everyone/anyone
- Other example: Satisfiability, given a boolean satisfiability formula,$(x_1 \vee x_2 \vee x_4), (\bar{x_2} \vee x_4 \vee x_1), \cdots$

$$\begin{aligned} \max \ & \# \text{ of clauses satisfied} \\ \text{s.t.} \quad & x \in \{0, 1\}^n \end{aligned}$$

Remember that optimal solution need not be unique. One of the special case is when variables have symmetry, in this case some kind of permutation can be applied to get multiple optimal solutions.

2 Examples

- Portfolio optimization – Every variable represents amount spend in each asset. Constraints might be on budget/availability/expected return. Objective is to minimize risk.
- Data fitting – The task is to find some model from some class of models which fit the data. What are the constraints, objective function, variables.

3 Classes of optimization problems

In general, we are interested in classes of optimization problems which can be solved easily or have specific properties. Different classes differ in what kind of constarints/objective functions are allowed for them. For example, Linear programming. (constraints and objective function has to be linear). We will study linear programming in detail later.

A class of problems is interesting if:

* Thanks to books from Boyd and Vandenberghe, Dantzig and Thapa, Papadimitriou and Steiglitz

– Many real world problems which can be modeled as the problem in the class
– Problems in the class are "easily" solved
– Problems in the class have nice properties (Duality), which can help us understand the structure. (more details later)

Our emphasis here is to learn some classes of optimization problem (Linear programming/ Semidefinite programming) and see how they can be applied to solve problems in computer science (complexity). There is a nice theory of how can we solve this problems (interior point, ellipsoid, simplex), but we won't care about that.

3.1 Least square problem

A least square problem is the one which does not have any constraint and the objective functions looks like $\sum_i (a_i^T x - b_i)^2$. Here $x \in \mathbb{R}^n$ is a variable with n coordinates. $a_i' s$ are vectors with n dimension and $b_i$are scalars.

$$\min \quad f_o(x) = ||Ax - b||_2^2 = \sum_i (a_i^T x - b_i)^2$$

Here A is a matrix of dimension $k \times n$. a_i^T are the rows of the matrix. One of the examples of such problem could be *Parameter estimation*, given a bunch of hyperplanes find the point closest to them.

This least square problem can be solved efficiently, when A is of full rank. We can prove that the best vector has to satisfy the equation $(A^T A)x = A^T b$.

Proof. Lets look at two spaces concerning A (matrix of dimension $m \times n$).

$R(A) = \{y \in \mathbb{R}^m :\sim Ax = y\}$(Range of A) and $N(A^T) = \{y \in \mathbb{R}^m :\sim A^T y = 0\}$(Null space of A^T).

Exercise: Show that $R(A) \oplus N(A^T) = \mathbb{R}^m$.

We can decompose $b = b_1 + b_2$,s.t., $b_1 \in R(A), b_2 \in N(A^T)$. Since $Ax - b_1 \in R(A)$, it is orthogonal to $b_2$for any x. The objective function can be written as $||Ax - b_1||^2 + ||b_2||^2$. So the optimal $x$satisfies, $Ax = b_1$. This implies $Ax = b - b_2 \Rightarrow (A^T A)x = A^T b$. □

Note: There is a nice way in which we can check if a particular problem is a least square problem or not. We will go through the basics later, which will enable us to prove this. Least square problems can be generalized too.

– Weighted least square problems:

$$\min \quad f_o(x) = ||Ax - b||_2^2 = \sum_i w_i (a_i^T x - b_i)^2$$

– Regularization:

$$\min \quad f_o(x) = ||Ax - b||_2^2 = \sum_i (a_i^T x - b_i)^2 + \rho \sum_i x_i^2$$

.